

REMARKS

Claims 20, 22-24 and 26-29 are pending in the application. Applicants herewith amend claims 20 and 26-29. Support for the Amendment is found at least at paragraph 42 of the published specification and in the claims as originally filed. No new matter is added. Entry of the Amendment is kindly requested.

I. Claim 20 is Proper

At page 2 of the Office Action, the Office objects to claim 20 because the limitation “proteins sequences” is allegedly “syntactically awkward.” To advance prosecution Applicants herewith amend claim 20 without prejudice or disclaimer. Applicants’ Amendment overcomes the objection.

Withdrawal of the objection is therefore kindly requested.

II. Claims 20, 22-24 and 26-29 Are Enabled Under 35 U.S.C. § 112, First Paragraph

At pages 2-5 of the Office Action, the Office rejects claims 20, 22-24 and 26-29 under 35 U.S.C. § 112, first paragraph, as allegedly lacking enablement because one skilled in the art would not know how to perform the recited method steps.

The Examiner must establish a reasonable basis to question the enablement of a claimed invention. *In re Wright*, 999 F.2d 1557, 1562, Fed. Cir. 1993). “It is incumbent upon the Patent Office, whenever a rejection on this basis [enablement] is made, to explain why it doubts the truth or accuracy of any statement in a supporting disclosure and to back up assertions of its own with acceptable evidence or reasoning which is inconsistent with the contested statement. Otherwise, there would be no need for the applicant to go to the trouble and expense of supporting his presumptively accurate disclosure.” *In re Marzocchi*, 439 F.2d at 224, 169 USPQ at 370.

Applicants respectfully disagree that a prima facie case of lack of enablement is set forth. At page 4 of the Office Action, the Examiner states that one of ordinary skill in the art would not know how to compare outlier protein sequences of a group of pathogenic organisms because the artisan does not understand “to what they are being compared.” To advance prosecution, Applicants herewith amend claim 20 to recite comparison of outlier protein sequences to

sequences present in public databases. Applicants' amendment overcomes this aspect of the rejection.

At page 4 of the Office Action, the Examiner further alleges that, "[m]erely comparing sequences would not necessarily lead one skilled in the art to [a] determination that the outlier proteins are unique." Claim 20 recites comparing the outlier protein sequences to the sequences within public protein sequence databases and identifying unique outlier proteins. The terms "unique," "homologous" and "identical" are well known in the art. For example, according to Fraser et al. (*Emerging Infectious Diseases*, Vol. 6, No. 5: 505-12 (2000)) "approximately one-quarter of the predicted coding sequences in each species are unique, with no appreciable sequence similarity to any other known protein sequences . . ." In addition, Koonin et al. (*Nature*, Vol. 420: 218-23 (2002)) state that "[o]ur current theoretical understanding of protein folding is insufficient to estimate the total possible number of proteins structures, but it too is likely to be vast. Obviously, only a miniscule fraction of the potential sequence space is populated by real protein sequences, but the number of unique sequences encoded in actual genomes is likely to be substantial . . ." Databases containing sequences are also well known in the art and are specifically set forth in the specification (e.g., at least at paragraphs 62 and 63 of the published specification).

Thus, one of ordinary skill in the art can readily and routinely obtain and use unique outlier protein sequences based on the comparison of a subject sequence to those sequences deposited in public databases.

The Examiner further asserts that it is possible that a unique outlier protein sequence could exist in nature (but not yet be deposited in a database) and thus not be "unique." Page 4, Office Action. The Examiner's position is contrary to law. Assuming *arguendo* that Applicants' method can possibly yield a false positive result, the inoperability of a single embodiment does not warrant a finding that the specification fails to enable Applicants' method claims under 35 U.S.C. § 112, first paragraph. The Court of Appeals for the Federal Circuit addressed this very issue of enablement when it stated that "[e]ven if some of the claimed combinations were inoperative, the claims are not necessarily invalid. 'It is not a function of the claims to specifically exclude...possible inoperative substances...' *Atlas Power Co. v. E.I. du Pont de*

Nemours & Co., 750 F.2d 169, 1576 (Fed. Cir. 1984). More to the point, “[t]he presence of inoperative embodiments within the scope of a claim does not necessarily render a claim nonenabled.” MPEP, 2164.08(b) 8th Ed. (August, 2005). Thus, the Examiner’s conclusion that Applicants’ method can yield a false positive result is insufficient to sustain a lack of enablement rejection.

At page 4 of the Office Action, the Examiner states, “or the unique outlier sequence may have matches to ESTs at any particular homology when in fact the sequence is truly unique to that particular pathogenic organism.” The Examiner’s argument is *reductio ad absurdum* because the Examiner assumes that ESTs can match polypeptides, which they cannot.¹

The Examiner cites to *In re Wands* to buttress the lack of enablement rejection, asserting that a large quantity of experimentation is necessary to determine if outlier proteins are unique to a pathogenic organism. The Examiner applies the wrong legal standard in making the rejection - the quantity of experimentation is not determinative but rather if the experimentation is routine in the art. The present facts resemble those under review in *In re Wands*, 858 F.2d 731, (Fed. Cir. 1988), wherein the Court reversed the Examiner’s rejection for lack of enablement holding that undue experimentation would not be required to practice the invention because it is known that in producing antibodies it is *routine* to first make monoclonal hybridomas to determine which hybridomas secrete antibodies with the desired characteristics. The Court found that the specification provided representative working examples (e.g., see Applicants’ Examples 1 to 7, pages 5 to 8 of the published specification and Fig. 1) as well as the methods (e.g., see Applicants’ published specification at paragraphs 71 to 89) needed to practice the invention. Like in the present case, Wands conducted successful experiments and demonstrated that at least one embodiment fell within the scope of the claims. The Court found the claims enabled based on the presence of Wand’s representative examples. Applicants examples are sufficient to enable one of ordinary skill in the art to make and use Applicants’ invention using only routine experimentation.

¹ Applicants refer the Office to *Just the Facts: A Basic Introduction to the Science Underlying NCBI Resources*, ESTs: GENE DISCOVERY MADE EASIER, available online at <http://www.ncbi.nlm.nih.gov/About/primer/est.html>,

Accordingly, withdrawal of the lack of enablement rejection is therefore kindly requested.

III. Claims 20, 22-24 and 26-29 Are Definite Under 35 U.S.C. § 112, Second Paragraph

Claims 20, 22-24, and 26-29 are rejected under 35 U.S.C. 112, second paragraph, as allegedly being indefinite.

A. At page 5 of the Office Action, the Examiner asserts that step (a) of claim 20, which recites “wherein said protein sequences are predicted from either whole genomic sequences, or from partial genomic sequences comprising at least one chromosome,” is allegedly indefinite. To advance prosecution, Applicants herewith amend claim 20 to recite “wherein said protein sequences are predicted either from whole or partial genomic sequences.” Applicants’ Amendment overcomes this aspect of the rejection.

B. At page 6 of the Office Action, the Examiner asserts that “comparing said outlier protein sequences of a group of pathogenic organisms” allegedly lacks antecedent basis in claim 20. To advance prosecution, Applicants herewith amend claim 20. Applicants’ Amendment overcomes this aspect of the rejection.

C. At page 6 of the Office Action, the Examiner states that “the databases” in claim 20 lacks antecedent basis. To advance prosecution, Applicants herewith amend claim 20. Applicants’ Amendment overcomes this aspect of the rejection.

D. At page 6 of the Office Action, the Examiner asserts that “the outlier protein” as recited in claims 26-28 allegedly lacks antecedent basis. To advance prosecution, Applicants herewith amend claim 20. Applicants’ Amendment overcomes this aspect of the rejection.

E. At page 7 of the Office Action, the Examiner asserts that the recitation of “selected in step (e)” in claim 28 is allegedly confusing because step (e), as amended, does not recite selection. To advance prosecution, Applicants herewith amend claim 20. Applicants’ Amendment overcomes this aspect of the rejection.

F. At page 7 of the Office Action, the Examiner asserts that claim 29 is allegedly indefinite because the structural element limitations of claim 29 do not limit the method of claim 20.²

Claim 20 recites calculating attributes from protein sequences wherein the attributes comprise specific biological parameters. Claim 29 reasonably apprises persons skilled in the art of the metes and bounds of the claimed invention (e.g., performance of an attribute calculation step). Claim 29 further limits the method of claim 20. Because one of ordinary skill in the art would understand that the attribute limitation of claim 29 is applied to the attribute limitation of claim 20 the rejection is improper.

In addition, the Office is respectfully reminded that to meet the goal of reaching a clearly defined issue for an early termination of proceedings, i.e., issuance of an Office Action or Allowance of claims, the Examiner is charged with conducting a careful and thorough search and fully applying the references in preparing the first Office Action on the merits in order for a speedy and just determination of the issues involved in the examination of the application. *See* MPEP §§ 706.07 and 904.03. The specification and claims were not duly considered prior to issuance of the outstanding Office Action. Applicants therefore kindly request that the rejection be withdrawn as premature.

Accordingly, withdrawal of the indefiniteness rejections is kindly requested.

In view of the above, reconsideration and allowance of this application are now believed to be in order, and such actions are hereby solicited. If any points remain in issue which the Examiner feels may be best resolved through a personal or telephone interview, the Examiner is kindly requested to contact the undersigned at the telephone number listed below.

² To clarify the basis of the rejection, Applicants held a telephonic interview with Examiner Moran on November 7, 2007 wherein the Examiner indicated that such was the interpretation of the language of the rejection.

The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

Respectfully submitted,



William J. Simmons, PhD.
Registration No. 59,887

SUGHRUE MION, PLLC
Telephone: (202) 293-7060
Facsimile: (202) 293-7860

WASHINGTON OFFICE

23373

CUSTOMER NUMBER

Date: November 26, 2007



A Science Primer

National Center for Biotechnology Information

About NCBI	NCBI at a Glance	A Science Primer	Databases and Tools
Human Genome Resources	Model Organisms Guide	Outreach and Education	News

About NCBI
Site Map

Science Primer:

Just the Facts: A Basic Introduction to the Science Underlying NCBI Resources

Bioinformatics

Genome Mapping

ESTs: GENE DISCOVERY MADE EASIER

Molecular Modeling

SNPs

Microarray
Technology

Molecular Genetics

Pharmacogenomics

Phylogenetics

Investigators are working diligently to sequence and assemble the genomes of various organisms, including the mouse and human, for a number of important reasons. Although important goals of any sequencing project may be to obtain a genomic sequence and identify a complete set of genes, the ultimate goal is to gain an understanding of when, where, and how a gene is turned on, a process commonly referred to as **gene expression**. Once we begin to understand where and how a gene is expressed under normal circumstances, we can then study what happens in an altered state, such as in disease. To accomplish the latter goal, however, researchers must identify and study the protein, or proteins, coded for by a gene.

As one can imagine, finding a gene that codes for a protein, or proteins, is not easy. Traditionally, scientists would start their search by defining a biological problem and developing a strategy for researching the problem. Oftentimes, a search of the scientific literature provided various clues about how to proceed. For example, other laboratories may have published data that established a link between a particular protein and a disease of interest. Researchers would then work to isolate that protein, determine its function, and locate the gene that coded for the protein. Alternatively, scientists could conduct what is referred to as **linkage studies** to determine the chromosomal location of a particular gene. Once the chromosomal location was determined, scientists would use biochemical methods to isolate the gene and its corresponding protein. Either way, these methods took a great deal of time—years in some cases—and yielded the location and description of only a small percentage of the genes found in the human genome.

Now, however, the time required to locate and fully describe a gene is rapidly decreasing, thanks to the development of, and access to, a technology used to generate what are called **Expressed Sequence Tags**, or **ESTs**. ESTs provide researchers with a quick and inexpensive route for discovering new genes, for obtaining data on gene expression and regulation, and for constructing genome maps. Today, researchers using ESTs to study the human genome find themselves riding the crest of a wave of scientific discovery the likes of which has never been seen before.

An Expressed Sequence Tag is a tiny portion of an entire gene that can be used to help identify unknown genes and to map their positions within a genome.

What Are ESTs and How Are They Made?

ESTs are small pieces of DNA sequence (usually 200 to 500 nucleotides long) that are generated by sequencing either one or both ends of an expressed gene. The idea is to sequence bits of DNA that represent genes expressed in certain cells, tissues, or organs from different organisms and use these "**tags**" to fish a gene out of a portion of chromosomal DNA by matching base pairs. The challenge associated with identifying genes from genomic sequences varies among organisms and is dependent upon genome size as well as the presence or absence of **introns**, the intervening DNA sequences interrupting the protein coding sequence of a gene.

Separating the Wheat from the Chaff: Using mRNA to Generate cDNA

Gene identification is very difficult in humans, because most of our genome is composed of introns interspersed with a relative few DNA coding sequences, or genes. These genes are expressed as proteins, a complex process composed of two main steps. Each gene (DNA) must be converted, or **transcribed**, into **messenger RNA** (mRNA), RNA that serves as a template for protein synthesis. The resulting mRNA then guides the synthesis of a protein through a process called **translation**. Interestingly, mRNAs in a cell do not contain sequences from the regions between genes, nor from the non-coding introns that are present within many genes. Therefore, isolating mRNA is key to finding expressed genes in the vast expanse of the human genome.

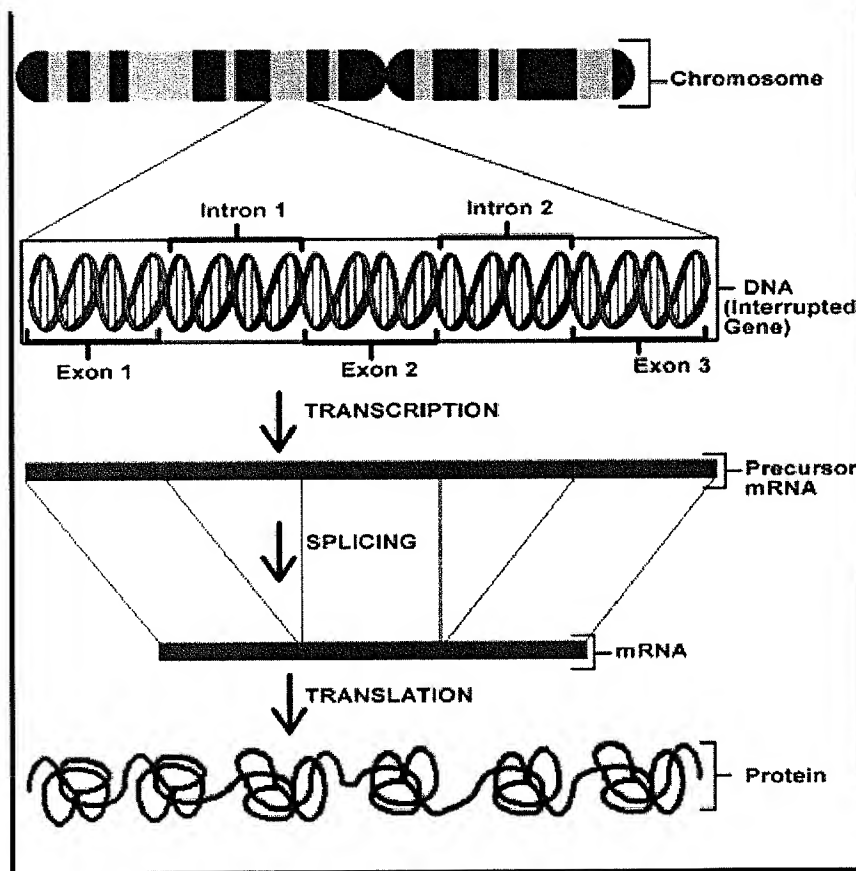


Figure 1. An overview of the process of protein synthesis.

Protein synthesis is the process whereby DNA codes for the production of amino acids and proteins. The process is divided into two parts: transcription and translation. During transcription, one strand of a DNA double helix is used as a template by mRNA polymerase to synthesize a mRNA. During this step, mRNA passes through various phases, including one called splicing, where the non-coding sequences are eliminated. In the next step, translation, the mRNA guides the synthesis of the protein by adding amino acids, one by one, as dictated by the DNA and represented by the mRNA.

The problem, however, is that mRNA is very unstable outside of a cell; therefore, scientists use special enzymes to convert it to **complementary DNA (cDNA)**. cDNA is a much more stable compound and, importantly, because it was generated from a mRNA in which the introns have been removed, cDNA represents only expressed DNA sequence.

cDNA is a form of DNA prepared in the laboratory using an enzyme called reverse transcriptase. cDNA production is

the reverse of the usual process of transcription in cells because the procedure uses mRNA as a template rather than DNA. Unlike genomic DNA, cDNA contains only expressed DNA sequences, or exons.

From cDNAs to ESTs

Once cDNA representing an expressed gene has been isolated, scientists can then sequence a few hundred nucleotides from either end of the molecule to create two different kinds of ESTs. Sequencing only the beginning portion of the cDNA produces what is called a **5' EST**. A 5' EST is obtained from the portion of a transcript that usually codes for a protein. These regions tend to be conserved across species and do not change much within a **gene family**. Sequencing the ending portion of the cDNA molecule produces what is called a **3' EST**. Because these ESTs are generated from the 3' end of a transcript, they are likely to fall within non-coding, or **untranslated regions (UTRs)**, and therefore tend to exhibit less cross-species conservation than do coding sequences.

A "**gene family**" is a group of closely related genes that produces similar protein products.

A UTR is that part of a gene that is not translated into protein.

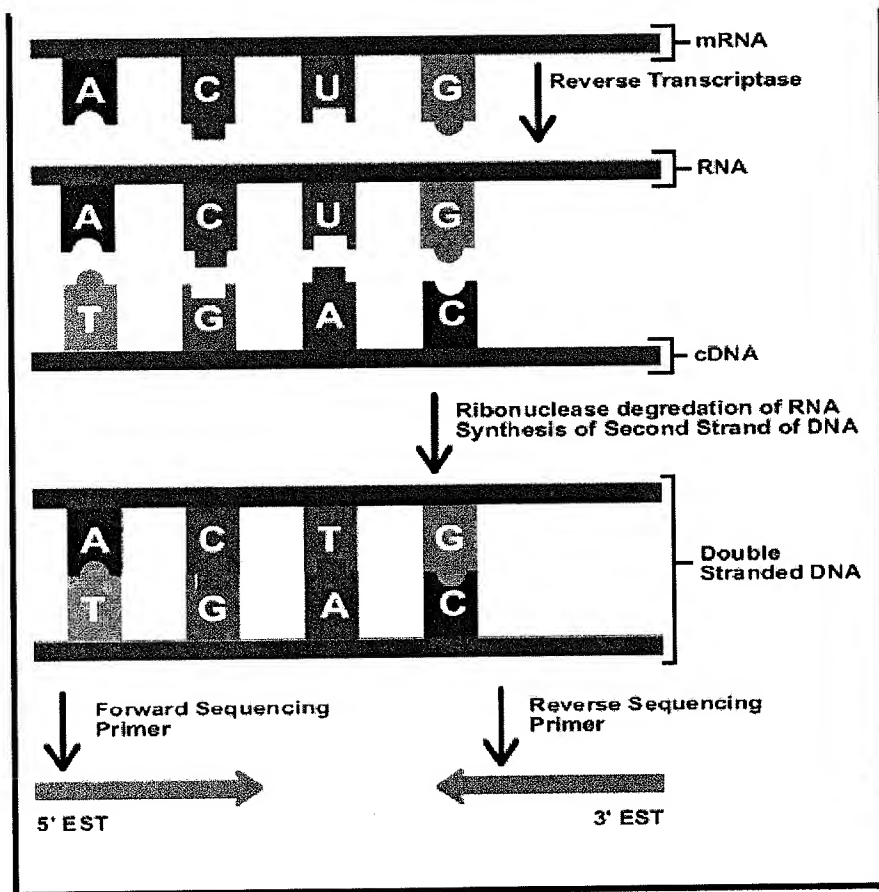


Figure 2. An overview of how ESTs are generated.

ESTs are generated by sequencing cDNA, which itself is synthesized from the mRNA molecules in a cell. The mRNAs in a cell are copies of the genes that are being expressed. mRNA does not contain sequences from the regions between genes, nor from the non-coding introns that are present within many interesting parts of the genome.

ESTs: Tools for Gene Mapping and Discovery

ESTs as Genome Landmarks

Just as a person driving a car may need a map to find a destination, scientists searching for genes also need **genome maps** to help them to navigate through the billions of nucleotides that make up the human genome. For a map to make navigational sense, it must include reliable landmarks or "markers". Currently, the most powerful mapping technique, and one that has been used to generate many genome maps, relies

on Sequence Tagged Site (STS) mapping. An STS is a short DNA sequence that is easily recognizable and occurs only once in a genome (or chromosome). The 3' ESTs serve as a common source of STSs because of their likelihood of being unique to a particular species and provide the additional feature of pointing directly to an expressed gene.

ESTs as Gene Discovery Resources

ESTs are powerful tools in the hunt for known genes because they greatly reduce the time required to locate a gene.

Because ESTs represent a copy of just the interesting part of a genome, that which is expressed, they have proven themselves again and again as powerful tools in the hunt for genes involved in hereditary diseases. ESTs also have a number of practical advantages in that their sequences can be generated rapidly and inexpensively, only one sequencing experiment is needed per each cDNA generated, and they do not have to be checked

for sequencing errors because mistakes do not prevent identification of the gene from which the EST was derived.

Using ESTs, scientists have rapidly isolated some of the genes involved in Alzheimer's disease and colon cancer.

To find a disease gene using this approach, scientists first use observable biological clues to identify ESTs that may correspond to disease gene candidates. Scientists then examine the DNA of disease patients for mutations in one or more of these candidate genes to confirm gene identity. Using this method, scientists have already isolated genes involved in Alzheimer's disease, colon cancer, and many other diseases. It is easy to see why ESTs will pave the way to new horizons in genetic research.

ESTs and NCBI

Because of their utility, speed with which they may be generated, and the low cost associated with this technology, many individual scientists as well as large genome sequencing centers have been generating hundreds of thousands of ESTs for public use. Once an EST was

For ESTs to be easily accessed and useful as gene discovery tools, they must be organized in a searchable

generated, scientists were submitting their tags to **GenBank**, the NIH sequence database operated by NCBI. With the rapid submission of so many ESTs, it became difficult to identify a sequence that had already been deposited in the database. It was becoming increasingly apparent to NCBI investigators that if ESTs were to be easily accessed and useful as gene discovery tools, they needed to be organized in a searchable database that also provided access to other genome data. Therefore, in 1992, scientists at NCBI developed a new database designed to serve as a collection point for ESTs. Once an EST that was submitted to GenBank had been screened and annotated, it was then deposited in this new database, called **dbEST**.

database that also provides access to genome data.

dbEST: A Descriptive Catalog of ESTs

Scientists at NCBI annotate EST records with text information regarding DNA and mRNA homologies.

Scientists at NCBI created dbEST to organize, store, and provide access to the great mass of public EST data that has already accumulated and that continues to grow daily. Using dbEST, a scientist can access not only data on human ESTs but information on ESTs from over 300 other organisms as well. Whenever possible, NCBI scientists annotate the EST record with any known information. For example, if an EST matches a DNA sequence that codes for a known gene with a known function, that gene's name and function are placed on the EST record. Annotating EST records allows public scientists to use dbEST as an avenue for gene discovery. By using a database search tool, such as NCBI's BLAST, any interested party can conduct sequence similarity searches against dbEST.

UniGene: A Non-Redundant Set of Gene-oriented Clusters

Because a gene can be expressed as mRNA many, many times, ESTs ultimately derived from this mRNA may be **redundant**. That is, there may be many identical, or similar, copies of the same EST. Such redundancy and overlap means that when someone searches dbEST for a particular EST, they may retrieve a long list of tags, many of which may represent the same gene. Searching through all of these identical ESTs can be very time consuming. To resolve the redundancy and overlap problem, NCBI investigators developed the **UniGene database**. UniGene automatically partitions GenBank sequences into a non-

redundant set of gene-oriented clusters.

Although it is widely recognized that the generation of ESTs constitutes an efficient strategy to identify genes, it is important to acknowledge that despite its advantages, there are several limitations associated with the EST approach. One is that it is very difficult to isolate mRNA from some tissues and cell types. This results in a paucity of data on certain genes that may only be found in these tissues or cell types.

Second is that important gene regulatory sequences may be found within an intron. Because ESTs are small segments of cDNA, generated from a mRNA in which the introns have been removed, much valuable information may be lost by focusing only on cDNA sequencing. Despite these limitations, ESTs continue to be invaluable in characterizing the human genome, as well as the genomes of other organisms. They have enabled the mapping of many genes to chromosomal sites and have also assisted in the discovery of many new genes.

[Back to top](#)

Revised: March 29, 2004

NCBI

NLM

NIH

[Privacy Statement](#)

[Disclaimer](#)

[Accessibility](#)